SICS/T-90/9001

Knowledge Acquisition Procedures for Diagnosis of Performance Problems
by

Anette Gäredal et al.

SICS technical report
T90001
ISSN 1100-3154

Knowledge Acquisition Procedures for Diagnosis of Performance Problems

Anette Gäredal,
Swedish Institute of Computer Science

David Melin-Högrell
Data Logic

Christer Nellborn
Svenska Institutet för Systemutveckling

Jan Olsson
IBM Lidingö

Abstract

This report describes knowledge acquisition work, performed during a three months period, as part of a knowledge-based system project at IBM. The aim with the work was to develop knowledge acquisition procedures for diagnosis of a specific performance problem.

1. Introduction

This report describes knowledge acquisition (KA) work, performed during a three months period, as part of a knowledge-based system project at IBM. The aim with the work was to develop knowledge acquisition procedures for diagnosis of a specific performance problem. The goal was to describe suitable procedures for extracting knowledge about how to diagnose a response time problem. The approach taken was empirical, i.e to develop a method(procedures) based on the results from actual knowledge acquisition sessions. Also the development of the procedures was based on a literature study which focused on KA techniques for diagnosis problems.

In the knowledge acquisition sessions different techniques (reported in various papers and books on knowledge acquisition), were used. These techniques and other techniques suitable for further knowledge acquisition sessions are described in this paper. Also a suggestion of different modeling and structuring techniques are presented. All the techniques are viewed and discussed with the particular problem domain in mind and no attempt is made to generalize the results for other areas or types of problems.

2. Techniques for Knowledge Acquisition

In KA work different techniques are needed for eliciting different types of knowledge. One of the first things to do during KA is to set the boundaries for the domain and the problems that are to be solved by the expert system. In this project these issues were tackled by using techniques listed as "techniques for domain organization" in the literature (see below) and based on the result, using some of these techniques, further refinement and KA was done.

2.1 Techniques for Domain Organization

There are mainly two different approaches for domain organization, one is to generate the concepts to be used in the system, the other is to organize and structure what has been generated. The approaches are (Boose, 1989):

"a) Generating concepts by:
- brainstorming (van Gundy, 1981)
- Crawford Slip method (Rusk and Krone, 1984)
- Q-methodology (McKeown & Thomas, 1988)

b) Organizing concepts into groups:
- expert identify common threads
- expert provide labels
- knowledge engineer records in the form of lists, notes, diagrams, charts, knowledge dictionary/glossary, (can use audiotape, videotape, whiteboard, wordprocessor, etc)
- Card sort, used for understanding how the expert conceptualizes the domain. Elements to be sorted comes from an analysis of the domain, a literature study or as a result from other elicitation techniques".

One of the first thing addressed when starting the elicitation was to decide the ambition level for the problem-solving capabilities of the KBS. This was tackled by letting the expert list the different problems types that could cause a response time problem. The resulting list was structured as a diagram and then further refined showing subproblems belonging to various problem types. The refinement consisted of:
- a provision of labels from the expert,
- identification of subproblems belonging to more than one problem type,
- recording of a knowledge glossary.

2.2 Techniques for Problem Solving Knowledge

The next thing to do was to acquire the knowledge that went in to diagnosing the response time problems. To acquire this knowledge other techniques were needed. Below is survey of some techniques for problem solving knowledge, extracted from (Diaper, 1989) and (Boose, 1989).

Interviewing (McGraw & Harbinson-Briggs, 1989)
used to identify tasks and major concepts, to structure and refine already acquired information.

Unstructured interview
To get a broad view of the domain. The unstructured interview technique is also referred to as: Overview interview, Grand Tour, see below. These interviews permit interviewees to cover topics in largely their on way. They provide "capacity for surprise" for

interviewers who have an idea of the kind of information that is needed and are prepared with a set of seed questions, prompts and probes.
A starter question in an unstructured interview could be: "Imagine you went into a bookshop and saw the book you wish you had had when you first started working in the field. What would it have had in it?" (Johnson, 1987a).
Another strategy might be: "When I read the documentation for this system, I had some trouble with (mention part or section). Can you explain to me?" (Davis, 1983),

The overview interview
The overview interview is done to get a general picture of the domain in order to understand how the expert thinks of his domain, what boundaries exist and how he sees them. The knowledge sought concerns the existence of subareas, different perspectives of the domain from the experts point of view, goals and organization of the domain. Example from the domain: Can you describe the things you do as a performance analysist? or When you visit a customer what are the things that you do?

Grand Tour (see also Knowledge Acquisition Grid below)
Aim to extract boundaries of the domain, different perspectives and goal organization. Example of question: "Please describe the kinds of things you do" or "Could you describe the kinds of things that schedulers do?"

Advantages:
- Appropriate when the knowledge engineer wants to explore an issue
- Facilitates description of domain in a way that is easy for the expert
- Goal is to establish rapport and get a broad view

Disadvantages:
- Data acquired is often unrelated and difficult to integrate
- Often exhibits lack of structure
- Does not allow gathering of specific knowledge
- Takes time and training to do well
- Similar questions asked in future sessions may annoy expert

Structured interview
Structured interview means an interview prepared in advance designed to elicit general knowledge of the domain. In the structured interview it is the knowledge engineer who guides the questioning according to the prepared questions. The same questions are asked in the same words and the same order for each interview. Questionnaires is a type of structured interview.
It is important to the design of structured interviews that the range of likely responses is adequately covered. If there exists a list of precoded answers the interviewer does not usually show the interviewee this list, nor are they usually read out, so it is important that the interviewer is familiar enough with the precoded answers to be consistent in categorizing the answers.
Example from the domain: When you talk about DASD problems are you considering several types of DASD problems?

Advantages:
- Forces an organization on the interview
- Very goal-directed
- Attempts to remove distortion from experts subjectively
- Allows better integration of material after the interview
- Forces the expert to be systematic
- Knowledge engineer identifies gaps in the knowledge which acts as a basis for questions
- Purpose of session is clear to expert

Disadvantages:

Needs more preparation by the knowledge engineer
Needs to study background material extensively

Focused interview
Focused interview is a specialized technique to elicit in a depth first fashion all that the expert knows and has to say about a particular single concept, operation, etc. The interviews can be broadly focused or narrowly focused. During a broadly focused interview a knowledge engineer may pose questions such as "describe the parts of a typical boiler system" whereas during a narrowly focused interview "are there any welding problems associated with carbon steel?" would be more typical.
Example from the domain: broadly - "describe vulnerable parts in the system"; narrowly - "can the size of CA & CI cause any problems when using VSAM as an access method?"

Structured or focused interviewing achieves results quickly and is appropriate when the knowledge engineer is fairly confident of his understanding of the domain. This understanding may result from prior knowledge or from the results of earlier elicitation sessions.

Case study
This technique focuses on the use of particular cases for eliciting knowledge. The expert reasons through the solution of a case, which can be of several kinds; typical , atypical, interesting or critical.
Four different case study techniques can be used:

Forward scenario simulation
The elicitor or the knowledge provider supplies an example situation or a case and the knowledge provider describes, usually step by step, what would be done in the situation or when handling the case. If the case is not a real one which the knowledge provider has processed, the technique suffers from the disadvantages of other hypothetical questioning strategies and thus an additional source of invalidity may be introduced. Also a warning against using this technique with knowledge providers who find it difficult to verbalize knowledge, a warning that should apply to any unstructured technique. We used this variant of case study technique in the problem domain.

Retrospective case description
This is similar to forward scenario simulation except that the cases are not hypothetical ones but are examples of ones which the knowledge provider has handled, preferably in the recent past.

Interesting cases
For some elicitation exercises the case chosen is of a specific type such as rare cases, those that take a particularly long time to solve, those which give the knowledge provider the most difficulty or are particularly memorable. There is a danger that in seeking examples of interesting cases the knowledge provider will have to delve too far into the past to be able to remember the case in detail.

Critical incidents
Memorable cases are the focus of the critical incident technique. As with studies of other interesting cases, it is a specialization of the retrospective case description. The knowledge provider is asked to recall and subsequently to discuss a specific incident or a case which was of critical importance in some context. the criticality of the

incident is thought to have strongly impressed details on the person's mind. The incident as a source for verbal reports is not only vivid, but in so far as the person relives the incident when it is brought to mind, it approximates a re-enactment of the incident which, it is thought, makes available kinds of knowledge which would not otherwise be recalled.

When it came to using any of the above listed techniques, the one chosen was the case study technique. It was considered to be the most appropriate to start with since cases with answers could be provided by one of the experts. Also this was a quick way of getting started without having to study the domain thoroughly, which is necessary to be able to pose any good questions. However, other techniques such as focused interviews, was considered for further elicitation.

Apart from the above mentioned techniques another interview technique used in the KA work rather extensively was "Teachback interviewing". It was mainly used for giving feedback to the expert, to retell a procedure or some information that had been elicited from the expert. It was not used in the same sense as described in the literature. Below is an extraction from (Johnson, 1987a) describing teachback interviewing:

Teachback interview

Teachback interviewing is based on the Conversation theory of Pask(1974), in which the notions of concepts and understanding are treated "as entities made public or objective by an interaction between participants". The technique is distinguished from other interviewing methods by being a participant activity rather then one between an interrogator and a respondent. It is very much "expert driven", in the sense that any topic to be discussed must originate from the expert, who has to specify at least one link with another topic.
The expert describes a procedure to the interviewer, who then "teaches" it back in the expert's terms and to the expert's satisfaction. When they agree, it is said that they share the same concept: this is Pask's Level 0. Level 0 answers are often explanations of how to do an algorithm. An example of a question that elicit the explication of a Level 0 procedure as an answer is: How do you do multiplication sums?
At Level 1 of the interaction, the interviewer asks the expert how the concepts was reconstructed, and the teachback procedure continues until the expert is satisfied with the interviewer's version. Then the interviewer has understood the expert. Level 1 is a matter of explanation of explanations; or perhaps, knowledge about knowledge (metaknowledge). It may correspond to the "epistemological level" identified by (Breuker, 1987) as "the missing level in developing experts systems"; at this level of analysis concepts are integrated into a framework of types of entities. At Level 1 a method for the reconstruction of Level 1 procedures is found. For example, how one remembers or recollects how to do a procedure is a typical Level 1 activity. Examples of questions that tend to elicit Level 1 activity are: How do you know multiplication always works? What do you need to know to be able to do it? How do you remember what to do? Do you know any equivalent methods? Why can you trust a method/system?

Relative to the above description of teachback interviewing, the technique was during the KA work used in a way that corresponds to Pask's Level 0. The interviewer taught back how to diagnose a certain response time problem. But it was not used any further, as described above for acquiring Level 1 procedures.

3. Description of KA procedures

This chapter describes a couple of steps to be taken and necessary preparations to be done during the KA work. The first thing to decide is the ambition level for the problem-solving capabilities of the KBS. This will delimit and guide the work to be done. Then it is necessary to choose the KA techniques to be used and make a rough schedule for how to proceed. Next comes preparations of the elicitation, what to read, which expert to use and so on. Also a summary of considerations to have in mind during the KA sessions is given. Thereafter some comments on how to process the material resulting from the KA sessions and finally a description of what needs to be included in the resulting conceptual model.

3.1. Ambition level

Deciding the ambition level for the problem-solving capabilities of the KBS is based on a number of aspects:

- how complex and difficult problems the KBS should be aimed at solving,
- (ideally), the proportion/severity of
  a) existing problem manifestations, and
  b) problems occurring in an average installation that the KBS will address,
- determining exactly what information will be made available to the KBS, and how

  it is structured. This is to ensure that none other is referred to during KA.

The above aspects are accomplished by using techniques for domain organization. When the domain has been roughly structured, the next step is to further define each problem area to a degree that makes it possible to determine how the KA for that area should best be done. This definition should probably be mostly verbal, but may also contain both quantitative and qualitative information.

The first step, once the problem areas have been defined and structured (e.g. as a hierarchical structure), is to select areas for further study. This selection should depend on factors such as

- frequency of problem
- severity of problem effects
- the problem's complexity
- availability of experts
- etc.

A good recommendation would be to start with a highly frequent problem that gives noticable effects. However the first couple of problems should not be too complex, it is better to work towards a growing complexity.

3.2. Selection of KA technique(s) for each problem area

Since there exist a multitude of techniques for eliciting problem-solving knowledge it is appropriate to choose only a limited number of techniques to start with. This should be done to be able to carry through the KA work with some (any) success. What KA technique can be considered appropriate for a given problem is dependent on several things, e.g.:

- the characteristics of the problem itself,
- the source of the knowledge,
- availability of the expert,
- knowledge already acquired.

This could be a difficult problem, at least for inexperienced knowledge engineers but also when it comes to "new" domains, i.e. types of problems that has not been reported on elsewhere in the research literature. In chapter 4 an approach for how to deal with this type of difficulty is described.

3.3. Preparations for Knowledge Acquisition

Regardless of what KA technique will be used, a number of things should be included in the preparations for a KA session. First of all, the purpose of the KA session (e.g. methodology development, problem area definition, problem manifestation classification, knowledge acquisition in general, KA for a specific area etc.) must be determined. But related to the purpose is of course the choice of KA technique.

Then set a date, coordinated with other activities, taking into account any natural sequence stemming from the structure of the domain or the chosen KA technique. Naturally, this is dependent on the availability of the experts and also on whether the same expert is to be used on several KA sessions, in which case it may be appropriate to group these together. Also, try to determine the approximate time necessary.

Thereafter select the Knowledge Engineer(s) that are to carry out the KA, and let the Knowledge Engineer(s) familiarize themselves with the subject, its place in the domain, and its terminology.
Determine what material (e.g. "case material" statistics, listings, manuals etc) will be needed for reference, and decide who should be responsible for preparing this. If appropriate, have the Knowledge Engineer(s) go through the material in advance.

Determine how much the expert should be told about the problem in advance (e.g. symptoms, installation characteristics, etc.). This information should approximate that which the expert usually has available when solving this type of problem.
Make sure that all aids (e.g. tape recorder, cassette tapes, pens, highlighters, paper clips, calculators etc.) are available and functioning.

3.4. The KA session

Unless there are excellent reasons for not doing this (e.g. that the KA consists of studying manuals), the KA session should be taped. All acronyms, abbreviations and technical terms that have not been previously used should be explained on the tape.

Everything referenced (e.g. in statistics listings, diagrams etc) should be marked with whatever is necessary to identify the KA session, and

- clearly identified on the tape
- highlighted (if possible, use different highlighting colors for different things)
- indicated, e.g. with a paper clip, in the margin of the list (to facilitate editing etc.)

Among the many things to remember during the KA session one of the most important is to listen carefully. For the answers to be of any value it is also important that the questions are well designed. There are a lot of advice available when it comes to the design of questions, (e.g. Moser and Kalton, 1978; CSO, 1976; Stacey, 1969). Among the things to avoid are:
- phrasing questions in the negative;
- making them so vague that there is both no suitable answer and many possible answers, e.g. "To what extent do you ...?";
- using jargon which is not usually used by the person;

- using pretentious language;
- phrasing questions so they are threatening or embarrassing;
- asking leading questions, sometimes called directive or loaded questions, i.e. ones which designate a particular response;
- asking double-barrelled questions , i.e. questions with two or more parts to them;

- asking questions which presume something about the person which may not be true, e.g. "How do you process ... [something the person may not process at all]?";

- asking hypothetical questions; and asking very long-winded questions.

Remember also that, for the purpose of KA, it is just as vital to find out why something is not of interest as why something else is, i.e. statements like "well, this looks pretty good so we needn't bother about this" should also be clarified.

3.5. Processing the results of the KA session

There are a couple of general things to do after each KA session both for documentation reasons and to facilitate the communication between the knowledge engineers.

An attempt to "translate" the result of a KA session into a brief verbal description (an "Abstract/Summary") of the expert's problem-solving should be done. Also add detailed references to all material (statistics, manuals) used by the expert. Refer to appropriate sections of the domain, i.e. the type of problem which is studied . Use a predetermined format for the Abstract/Summary, that should include:

- Date
- Expert
- Knowledge Engineer(s)
- Purpose (e.g. "overview of problem area")
- Problem Area (e.g. "MRO definitions")
- KA technique used
- For each specific technique list any prerequisite; e.g. for a case study:
  - installation characteristics (hardware and software)
  - materials used for reference (e.g. statistic type)
  - variables in the reference data (e.g. "CPU utilization")
  - problem indications (e.g. "gradual response time degradation for DL/I transactions")

- Diagnosis (e.g. "too few threads defined")
- Recommendations about how to correct the problem (e.g. "buy a bigger CPU"), and, for each recommendation: punch, possible trade-offs (e.g. new problems that the recommendation may cause)
- General comments and impressions

Have the tape of the KA session transcribed immediately after the interview, and use a word processor for transcription. The transcript should then be edited by one of the Knowledge Engineers, using a common format (reference numbers pointing to the materials used, and comments and/or clarifications from the notes should be added to the transcript, but in a different typeface from the actual interview).

Everything from the KA of a specific problem area should be filed in an orderly manner, including:
- tape(s)
- reference materials used
- transcripts
- notes

• diagrams of model
All acronyms, abbreviations and technical terms should be added to a knowledge glossary.This is also the time when all material (lists etc) referenced during the KA should be processed (e.g. indexed).
Further processing of transcripts is done by using mediating representations to describe the knowledge acquired during the KA session (see chap. 5).

3.6. Conceptual model

To facilitate development of a conceptual model, it is necessary to first establish a common ground, i.e. defining the concepts relating to the model. These concepts can be divided into three classes, Domain concepts, KA concepts and Knowledge Representation (KR) concepts. It should be noted, however, that many concepts belong in more than one class, and that this possibly even is a prerequisite for a smooth development of the KBS. The concept definitions are then set down in a "knowledge glossary".

Domain concepts
The domain concepts are those used by the expert(s) and refer to the domain itself. These definitions can usually be obtained from textbooks or manuals, or through interviewing domain experts (who may disagree on some definitions...).
Depending on the level of ambition, the domain model can be more or less detailed, showing all or just a selected number of concepts and connections between these. One possible way of doing this is to describe a "network model" of the major concept areas, decomposing each area into its components, and then describing "sub-networks" for each area.
It should be noted that there is no single objective and independent way of modeling a domain. The model will always reflect the purpose ("problem-solving", "descriptive", etc.) of those constructing it. Examples of domain concepts are: "MAXTASK", "MRO", "DASD".

KA concepts
This class includes concepts used in Knowledge Engineering, and refers primarily to the expert's problem-solving process, both from the viewpoint of the domain expert and from that of the end-user of the KBS. These are knowledge structuring concepts and should preferably be independent of the final KR/KB structure. An example of a KA concepts used is: "Hypothesis".

KR concepts
The final class contains all concepts that are dependent on the Expert System shell and the final KR/KB structure.
Examples of KR concepts used are: "Objective", "Explanation set", "Evidence" etc.

Both the Domain and the KA concepts are used in the modeling process which results in a conceptual model. During the modeling process different mediating representations are chosen, for describing the knowledge as identified by the Domain and KA concepts.

The KR concepts are used when translating the conceptual model to the final implementation.

Two types of interdependencies can be defined, within or between the concept "classes" described above. The dependencies within each class can be represented in an entity/relationship diagram, while those between concepts in different classes can be described in a matrix (or even through simple cross-references in the definitions). Examples of concepts, used in the domain, with connections between them are: "Problem

Areas", "Problem Manifestations", "Information needed (to solve the problem)" and "Recommendation".

The knowledge to be acquired can be said to be of two types:
a) knowledge to be put into the Knowledge Base and
b) knowledge about the domain itself and its characteristics, used in the KA process.

Both types affect the structure of the Knowledge Base, as well as the design of the Expert System as a whole. Examples of knowledge to be acquired in the domain of response time problems are:

- the domain structure, i.e. how to break down the domain into separate well-defined

  problem areas (based on diagnoses),
- how the problems can be categorized, i.e. the different problem manifestations that appear to the user of the system,
- the possible diagnoses for a given problem manifestation,
- the problem-solving process used by the expert(s) whose knowledge we want to use.

  Expressed in terms of the concepts relating to the conceptual model and then translated to the knowledge representation,
- the information needed to solve a problem, how it is structured, and where it can be

  found (i.e in log's, measurement data from the system, etc.),
- the recommendation(s) regarding suggested remedies for the problem, once it has been

  diagnosed,
- the trade-offs for a given recommendation,
- which way of presenting the results arrived at by the KBS is most appropriate to the way the user works.

4. A combination of different elicitation techniques: Knowledge Acquisition Grid

One approach that was examined, through a literature study, during the three-months period was the Knowledge Acquisition Grid, reported in (LaFrance, 1988). This approach combines different types of knowledge with different types of questions and gives suggestions when to use what. By using the suggested matrix, (see fig. 4.1), and making entries in it throughout the KA work the knowledge engineer can get a good view of what has been covered and what remains to be done. The matrix is a helpful aid for an inexperienced knowledge engineer. Below is a description of the KA Grid, extracted from (LaFrance, 1988):

The "Types of Questions" dimension of the KA Grid (LaFrance, 1988), articulates six distinct kinds of questions directed at different aspects of expert knowledge. The KA Grid organizes expert knowledge and knowledge engineer questions as separate but interacting dimensions. The first dimension describes the "Forms of Knowledge" in which an expert's know-how is stored. The second dimension describes the "Types of Questions" that are available to knowledge engineers directed to making the know-how explicit.

	Forms of knowledge				
Question Types	Layouts	Stories	Scripts	Metaphors	Rules-of -thumb
Grand Tour					
Cataloging Categories					
Ascertaining Attributes					
Determining Inter-connections					
Seeking Advice					
Cross-checking					

Figure 4.1. The Knowledge Acquisition Grid

FORMS OF KNOWLEDGE:

Layouts
Incorporates the expert's "map" of the task, including an understanding of its boundaries, organization, and basic classifications. Layouts subsume and give coherence to the expert's facts and heuristics by specifying the goals to which they aim, and the criteria used to characterize the problem at hand. By getting access to how the expert sets the task and organizes current information in light of prior knowledge and the present context, the knowledge engineer is better able to understand how to frame the problem.

Stories
Represents the classic cases and typical examples culled from the expert's long experience with the problem domain. Stories can be of a number of types in addition to the now familiar *Talking Aloud Protocol* (Waterman, 1971). For example *Explanatory Stories* are

those in which the expert, seeking to account for some puzzling situation, narrates a set of events that unfold in such a way as to lead up to and produce the phenomenon in question. A *Diagnostic Story/Prescriptive Story* describes some phenomenon in a way that shows what was wrong and what needed fixing.

Scripts
Scripts give the expert's sequential and procedural knowledge of the domain. The basic elements of a script are its roles, standard props or objects needed to carry out the actions, as a standard sequence of scenes wherein one act enables the next, and some normal results from performing the activity successfully (Abelson,1981). To know an expert's scripts is to have a temporal chart of critical actions, and to be able to understand each action in terms of the prior knowledge required to perform it.

Metaphors
Encapsulates the expert's alternative images of the task, each of which includes unique features, constraints, and options. Metaphors describe one thing by reference to another apparently dissimilar thing so that the first is understood more completely than if the comparison had not been made. Their advantage is being able to present an idea which can later be reconstructed and embellished through probes directed to the expert by the knowledge engineer.

Rules-of-thumb
Provides the myriad tactics and heuristics for interpreting and dealing with the array of circumstances encountered in carrying out the task. A rule-of-thumb encapsulates tacit knowledge about which conditions warrant which actions, and about how to gather data on and assess current conditions. Rules-of-thumb are concrete, implementable strategies of minor to moderate scope which can single out and define as issues those specific, limited conditions for which they serve as the most complete strategy.

TYPES OF QUESTIONS:

Grand Tour
Aim to extract boundaries of the domain, different perspectives and goal organization. Sub-areas within the expert's domain are comparably pursued by means of *Mini-Tour* queries. Example of question: "Please describe the kinds of things you do" or "Could you describe the kinds of things that schedulers do?"

Cataloging the categories
The expected outcome of these questions is an organized taxonomy of the expert's terms and concepts. An example of a Cataloging the Categories Question is: "When you gave me an overview of your job, you talked about schedulers. Are there different types of schedulers? Are schedulers a subtype of some other kind of job?" This type of question might follow a *Layout* response to a *Grand Tour Question.*

Ascertaining the attributes
These questions aim to discover the distinguishing features and range of possible values of the expert's concepts. An example of an ascertaining the attributes question is: "You've described a number of types of scheduling situations that you've encountered. I wonder whether you could now take the first two that you mentioned, and describe some ways in which these two are similar to each other but different from the third example that you gave". As stated, this particular question is a follow-up to a *Stories* description.

Determining the interconnections
These questions directed at uncovering the relations among the concepts in the domain. Of particular interest is the existence of a causal model for the whole domain or parts of it. An example of Determining the Interconnections Question is: "In describing the routine set of steps for scheduling an order you said that checking the request date occurs before anything else. Why is that the case?" This specific question is directed at previously obtained *Scripts* information.

Seeking advice
These questions are designed to reveal the expert's recommendations and hence strategies for how to deal with a variety of conditions such as how to determine current conditions and which conditions warrant which actions. An example of a Seeking Advice Question would be: "You've compared scheduling to playing a board game; from your experience with playing board games what advice could you give doing scheduling?" This type of question would be in response to a previously volunteered *Metaphor*.

Cross-checking questions
These questions are designed to validate and examine the limits an previously obtained information. Cross-checking questions actually consist of five subtypes including, the *Naive Question*, *Playing Devil's Advocate*, *Posing Hypothetical Situations*, asking *How sure are you?* and *Seeking the Exception.* An example of a Naive Question is "Bear with me while I ask what may appear to be a naive question. Could you tell me why orders need to be scheduled?" An example of a Devil's Advocate Question would be: "Let me play devil's advocate in response to your story about the need to set priorities on scheduling multiple orders. What if you didn't set priorities?"

One thing that remains to be included in an approach such as the KA Grid is how to process and model the material resulting from use of the Grid. In (LaFrance, 1988) no such suggestions or recommendations are given. During the period which this report covers, an attempt to come up with suggestions of modeling and processing techniques was made and they are described in the following chapter.

5. Processing and Modeling

It is important that the result from each KA session is processed and documented in some way. Depending on what kind of elicitation technique is used and what kind of knowledge is at hand, different ways of processing and modeling the material can be used. In this chapter a summary (with examples) for each of the suggested approaches, which are used as mediating representations, is given.

5.1 Processing data for Domain Organization

Inherent in some of the domain organization techniques are a processing flavor, i.e. the techniques in themselves aim at organizing and structuring the domain. The resulting structure can then be used for the processing of problem-solving data.

An example: During the domain organization a description of different categories in the domain is elicited. Such as
- indication,
- recommendation,
- etc.

These categories can then be used when processing material from for example a case study (a more detailed example is given below, see "defined framework").

5.2 Modeling Problem Solving knowledge

Several different approaches when modeling problem solving knowledge is necessary, both when it comes to processing material resulting from the use of different elicitation techniques and for gathering different types of knowledge. The following survey describes different approaches that has been used during the KA work.

• A defined *framework* for gathering descriptive knowledge about each identified problem area. In the recorded (audio, video, protocol transcript, etc.) material look for:

DEFINITION: of the problem,
CAUSE: what is/are the cause/s for this problem,
INDICATION: in what way do the cause show up,
RECOMMENDATION: how can the cause be handled,
TRADE-OFFS: what can the recommendation affect,
REPORTS: in what report/s (generated by the system) is/are the indication/s listed.

With all material processed this way a uniform representation and a possibility to group together concepts is created.

• A formalized grammar - *Systemic Grammar Network* (SGN) (see Johnson,1987), for a further characterization and specification of the domain. In this case it can be used for describing the various threshold levels to be used in the system , see fig.3.1. A mutually exclusive choice is denoted by a straight bracket and a co-occurring choice by a curly bracket. (There is also a recursion symbol and a conditional link.)

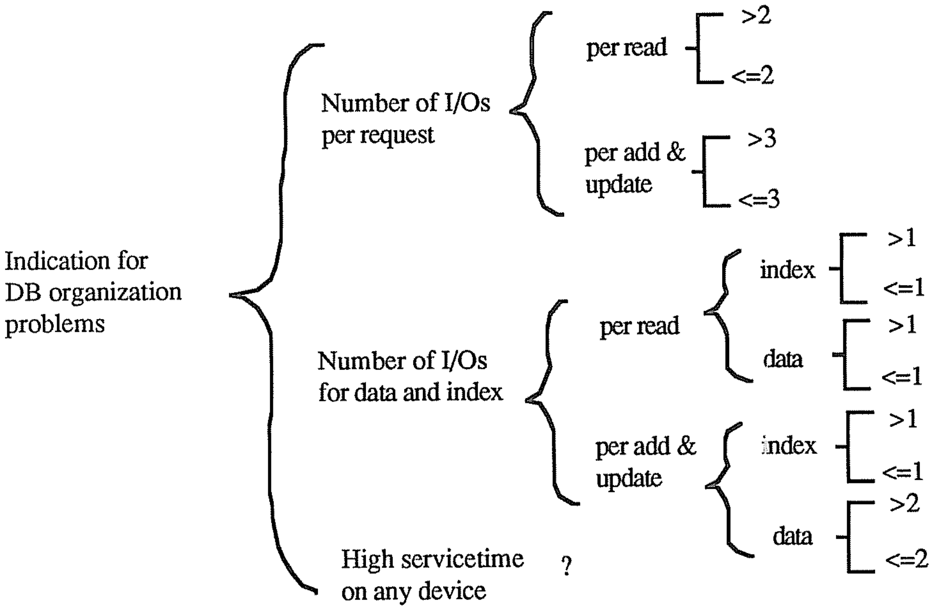

Figure 5.1. Example using SGN

By this type of representation it becomes obvious what parts are missing or what parts need to be further elaborated.

• A *"Step-by-Step" description* to enable capturing the reasoning process. Analyze a protocol transcript and look for: Input, Action, Results and Conclusion both intermediate and final. Identify also the various key concepts that appears under each label (underlined in the example).
Example:

ACTION: Analyze the Direct Access Device Activity report. Check the average response time. Should be below 25ms - 30ms for all devices.
RESULT: Average response time low.
CONCLUSION: No problem with I/O subsystem.

The above representation is an easy way to describe an ongoing process and it can be used for extracting problem-solving strategies from protocol transcripts.

• *ABC-modeling* (report forthcoming) to further illustrate the reasoning process, see fig.3.2. The "Step-by-Step" description is a good input to this modeling approach.

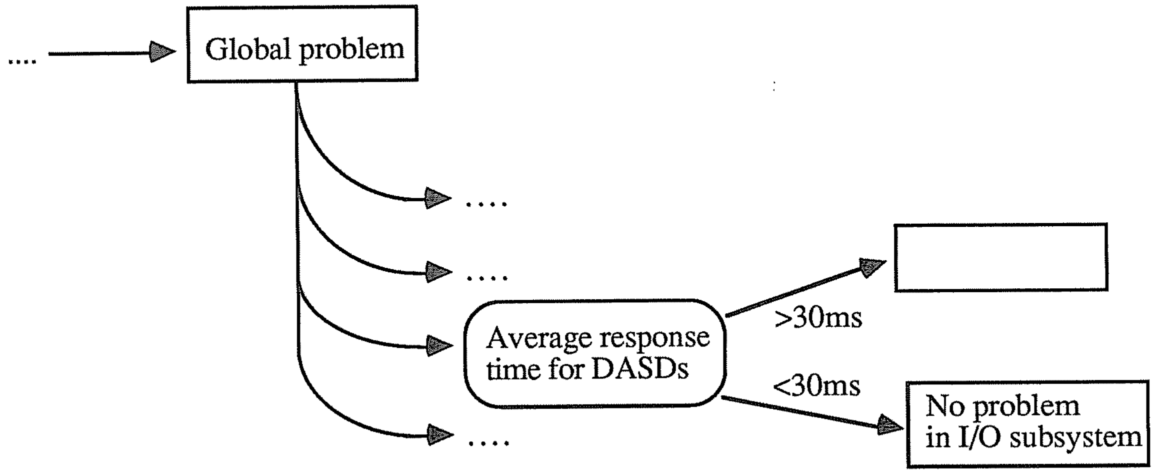

Figure 5.2. Example using ABC-modeling.

In the above representation each symbol denotes a special meaning. By using this approach it is possible to get an overview of how much has been covered of the problem-solving process.

• *Cause - Effect relations* (and others) to illustrate the complexity and connection between different problem areas, see fig. 3.3.

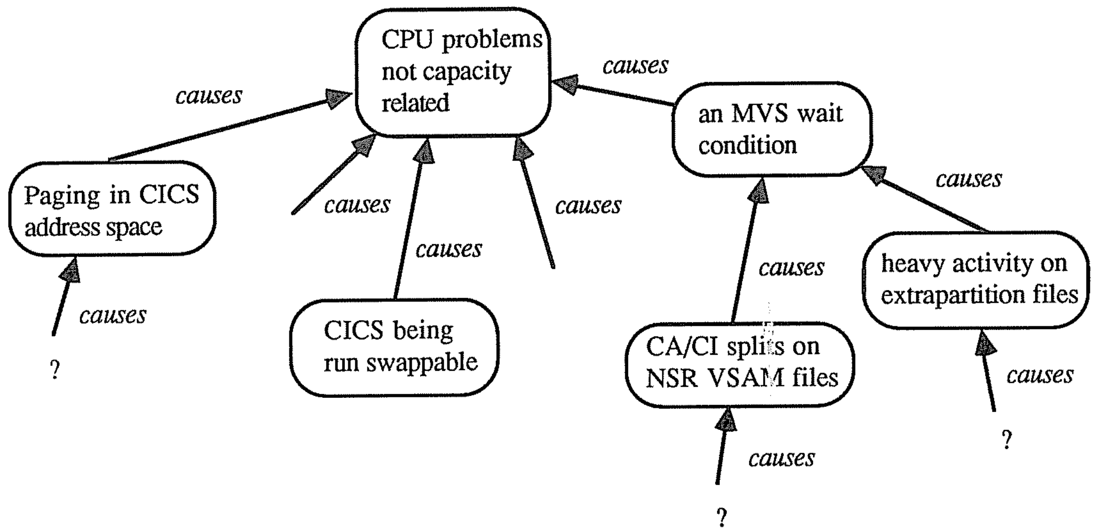

Figure 5.3. Example using Cause-effect relations.

This type of network gives a view of what problem areas are related and also in what order different problems occur.

• *If -Then descriptions* to capture heuristics and rules-of-thumb.

IF the CPU usage is over 85%
THEN it is suspect.

Each of these modeling approaches points to holes, misconceptions, inconsistencies, etc. in the elicited material. Modeling is an iterative process and the modeled material should lead to new KA sessions until the expert and knowledge engineer agrees on a final version of the knowledge base. Often very specific questions arises during the modeling and both structured and focused techniques can be used when continuing the elicitation.

6. Summary

During the KA work different "stages" of knowledge were identified, they were:

a) The "raw" knowledge, e.g. that acquired from a domain expert through the use of elicitation techniques,
b) The "structured" knowledge, e.g. an edited case study abstract,
c) The "modeled" knowledge, described through the use of a mediating representation in the conceptual model,
d) The "translated" knowledge, i.e. described in the terms and structure of the Knowledge Representation, e.g. input through a "KB Editor" which describes the equivalence between the "modeled" and "translated" knowledge and
e) The "production" knowledge, in the format stored in the Knowledge Base and used by the shell.

To accomplish the transition between the different stages some formalized procedures are needed. During the three-months period guidelines for some of these transitions were developed. For the transition from "raw" to "structured" knowledge, advice about how to construct an abstract/summary was given. Also, in chap. 3, recommendations was given about the resulting documentation. For the transition from "structured" to "modeled" knowledge a number of modeling approaches are described in chap. 5. The transition between the other two stages remains to be developed.
In this work a need for guidelines concerning the relation between elicitation technique and modeling approach became obvious and below is an attempt in giving such guidelines.

6.1 Relation between modeling approach and elicitation technique

Most of the modeling approaches described above (see chap. 5) can be used on material from any session using whatever elicitation technique. However, the modeling approach chosen might be more or less suited for the knowledge in the material. Also the elicitation techniques might be more or less suited for extracting the knowledge one wants to acquire. Below is a summary relating the various elicitation techniques and modeling approaches. First a brief overview of what each elicitation technique and modeling approach aimed at in terms of knowledge types.

ELICITATION TECHNIQUES:
Interview - used to identify tasks and major concepts, to structure and refine already-acquired information.
Unstructured interview - get a broad view of the domain.
Structured interview - elicit general knowledge of the domain.
Focused interview - all about a particular single concept, operation, etc.
Case Study - reasons through the solution of a case
Knowledge Acquisition Grid - articulates six distinct kinds of questions directed at different aspects of expert knowledge.

MODELING APPROACHES:
a defined *framework* - models descriptive knowledge,
Systemic Grammar Network (SGN) - gives a characterization and a specification of the domain,
Step-by-Step" description - aims at capturing the reasoning process,
ABC-modeling - further illustrates the reasoning process,
Cause - Effect relations (and others) - shows relations between areas, objects, etc.
If -Then descriptions - describes heuristics and rules-of-thumb.

When using the *interviewing* technique; either the unstructured, structured or focused one, the modeling approaches to use are the *frame-work* and the *SGN* approach, possibly could also the *Cause - Effect relation* approach be of use if the interview has covered any such aspects.

For *Case-studies* probably the best way of modeling the resulting transcript to acquire the reasoning process is to use the *Step-by-Step* description first and then the *ABC-modeling*. Other valuable information that might be included in the transcript such as threshold levels and/or rules-of-thumb can be captured by *If -Then descriptions* for the latter and *SGN* for the former.

	Modeling approaches					
Elicitation techniques	Framework	SGN	Step-by-Step	ABC-modeling	Cause-Effect relations	If-Then
Unstructured interview	X	X			(X)	X
Structured interview	X	X			(X)	X
Focused interview	X	X			(X)	X
Case Study		X	X	X		X
Knowledge Acquisition Grid	X	X	X	X	X	X

Figure 6.1. Combinations of elicitation techniques and modeling approaches.

In the matrix (fig. 6.1.) there is no ordering of the modeling approaches for any of the elicitation techniques. Instead each modeling approach suitable for any elicitation technique is marked with an X. The interview technique "teachback interview" has not been included in the matrix since it was mainly used for feedback.
The suggested matrix is based on the use of elicitation techniques and modeling approaches used during the three-months period described in this report.

7. References

(Abelson,1981) Abelson, R. P., *The psychological status of the script concept*, American Psychologist 36, 71-729, 1981.

(Boose, 1989) Tutorial on Knowledge Acquisition, IJCAI'89, Boose,J. and M. Shaw, 1989.

(Breuker, 1987) Breuker, J. A. and B. J., Wielinga, *Use of models in the interpretation of verbal data*, Kidd, A. (ed), Knowledge acquisition for expert systems: a practical handbook, London: Plenum Press, 1987.

(CSO, 1976) *Ask a silly question,* booklet prepared by the Survey Control Unit of the Central Statistical Office, London: HMSO, 1976.

(Davis, 1983) Davis, W. S., *Tools and techniques for structured systems, analysis and design*, London: Addison-Wesley, 1983.

(Diaper, 1989) Diaper, Dan (ed.), *Knowledge elicitation; principles, techniques and applications*, Ellis Horwood Limited, 1989.

(Johnson, 1987a) Johnson, L. and N. E., Johnson, *Knowledge elicitation involving teachback interviewing*, Kidd, A. L. (ed) Knowledge acquisition for expert systems: a practical handbook, London: Plenum Press, 1987.

(Johnson, 1987) Johnson, N.E, *Mediating Representations in Knowledge Elicitation*, In Proceedings of the first European Workshop on Knowledge Acquisition for Knowledge-Based Systems, Reading University, 1987.

(LaFrance, 1988) LaFrance, Marianne. *The Knowledge Acquisition Grid: a method for training knowledge engineers*, Boose J. & B. Gaines (eds.), Knowledge-Based systems Vol.1., Academic Press, 1988.

(McGraw , 1989) McGraw, L. and K. Harbinson-Briggs, *Knowledge Acquisition - principles and guidelines*, Prentice Hall, 1989.

(McKeown & Thomas, 1988)

(Moser, 1978) Moser, C. A. and Kalton, G., *Question wording,* Brynner, J and Stribley, K. M. (eds) Social research: principles and procedures, London: Longman in Association with the Open University, 1978.

(Pask, 1974) Pask, G., *Conversation, cognition and learning: A cybernetic theory and methodology*, London: Elsevier, 1974.

(Rusk, 1984) Rusk, R. A. and Krone, R. M., *The Crawford Slip Method (CSM) as a tool for extraction of Expert Knowledge*, In G.Salvendy (ed.), Human-Computer Interaction, New York: Elsevier, p. 279-282, 1984.

(Stacey, 1969) Stacey, M., *Methods of social research*, Oxford: Pergamon Press, 1969.

(Waterman, 1971) Waterman, D. A. and A. Newell, *Protocol analysis as a task for artificial intelligence,* Artificial Intelligence, 2, p. 285-318, 1971.